International Journal of Reconfigurable and Embedded Systems (IJRES)
Vol. 11, No. 2, July 2022, pp. 166~174
ISSN: 2089-4864, DOI: 10.11591/ijres.v11.i2.pp166-174

Comparison of breast cancer classification models on Wisconsin dataset

Rania R. Kadhim, Mohammed Y. Kamil
College of Sciences, Mustansiriyah University, Baghdad, Iraq

Article Info

Article history:

Received Feb 5, 2022
Revised Feb 21, 2022
Accepted Mar 11, 2022

Keywords:

Breast cancer
Classification
Computer-aided diagnosis
Machine learning
Wisconsin

ABSTRACT

Breast cancer is the leading cause of death for women worldwide. Cancer can be discovered early, lowering the rate of death. Machine learning techniques are a hot field of research, and they have been shown to be helpful in cancer prediction and early detection. The primary purpose of this research is to identify which machine learning algorithms are the most successful in predicting and diagnosing breast cancer, according to five criteria: specificity, sensitivity, precision, accuracy, and F1 score. The project is finished in the Anaconda environment, which uses Python's NumPy and SciPy numerical and scientific libraries as well as matplotlib and Pandas. In this study, the Wisconsin diagnostic breast cancer dataset was used to evaluate eleven machine learning classifiers: decision tree, quadratic discriminant analysis, AdaBoost, Bagging meta estimator, Extra randomized trees, Gaussian process classifier, Ridge, Gaussian nave Bayes, k-Nearest neighbors, multilayer perceptron, and support vector classifier. During performance analysis, extremely randomized trees outperformed all other classifiers with an F1-score of 96.77% after data collection and data analysis.

Corresponding Author:

Mohammed Y. Kamil
College of Sciences, Mustansiriyah University
St. Palastine, Baghdad, Iraq
Email: m80y98@uomustansiriyah.edu.iq

1. INTRODUCTION

Breast cancer is the second major cause of death in women, coming in second only to lung cancer [1]. Cancerous breast cells multiply unrestrained, which is one of the most telling signs that something is wrong with your breasts. Until today, it has been impossible to stop the growth of breast cancer. Breast tissue tumors should be found as early as possible to maximize the chances of survival, as this disease accounts for roughly 15% of all cancer-related fatalities [2]. Tissue biopsies can be used to detect cancer in its earliest stages, increasing the possibility of a positive result. It is possible to identify this disease using various methods, including a biopsy, ultrasound, thermography, and fine-needle aspiration biopsies [3]. While mammography has been the primary mode of early detection, it is not always enough for doctors to conclude whether or not the patient has breast cancer [4], [5]. As a result, the detection rate is just (60-70)% accurate [6]. The patient will need to take additional testing, which is expensive and time-consuming [7].

A biopsy is the only effective method of determining if a woman has breast cancer [8]. A biopsy sample is obtained using specialized needle equipment using an imaging test such as X-rays or other imaging methods. A small metal marker may be inserted into the breast to facilitate future imaging examinations. Scientists examine the cells in a lab to determine if they are malignant. It is also critical to know what type of cells are involved in breast cancer, how aggressive it is (grade), and if the cells have hormones or other

receptors that could modify your therapy options [9]. Breast biopsy results can take a few days to process. Physicians are experts in the study of blood and bodily tissues and perform their work in a specialized laboratory after the biopsy is completed [10]. The pathologist's report details the size and consistency of the tissue samples and their location, whether they are malignant or not. As a result of their differing opinions, these experts may need more surgery to collect additional tissue for location study [11]. Conventional diagnostic procedures have limitations researchers have applied machine learning (ML) methodologies to provide a second opinion to clinicians and reduce the risk of human error, resulting in patient death [12]-[15].

ML approaches help automate the decision-making process and help increase a decision support system [16]. An increase in the amount of work that doctors must accomplish is made possible by better precision and speed of response, especially in times of medical staff shortage. Preventive care and individual mental health therapies can benefit from decision-support and health-monitoring systems that apply ML techniques to learning [17]. Using ML classifiers and the Wisconsin diagnostic breast cancer (WDBC) dataset, this research evaluated approaches for detecting breast cancer.

Salama *et al*. [18] presented a comparison of the decision tree (J48), multi-layer perception (MLP), sequential minimal optimization (SMO), naive Bayes (NB), and instance-based for K-nearest neighbor (IBK). The results indicate that classification using SMO alone or in combination with MLP or IBK is superior to other classifiers. SMO outperforms other classifiers in terms of accuracy 97.7%. Azar *et al*. [19] implemented an ANN algorithm that included multilayer perceptron (MLP), probabilistic neural networks (PNN), and radial basis function (RBF). The testing and training phases determined that PNN was the high accurate classifier, with accuracy, sensitivity, specificity, precision, and area under the curve (AUC) of 98.66%, 95.65%, 95.82%, 97.77%, 0.99, respectively. Senapati *et al*. [20] proposed the local linear wavelet neural networks. The network was trained to detect breast cancer in the recursive least squares (RLS) technique to parameters. Discrepancies in the local linearity network of wavelet neurons with a traditional wavelet neural architecture (WNN) refer to the weighted average of the connections between the two conventional WNN's input, and output layers are replaced using a linear regression model. The high accuracy value obtained was 97.2%. Dora *et al*. [21] proposed a unique Gauss-Newton representation-based algorithm (GNRBA) for classification. It uses sparse representation in conjunction with training sample selection. Compared to the traditional l1-norm approach, this method evaluates sparsity more computationally efficiently. The suggested method achieved the greatest classification accuracy of 98.25%, according to the results. For 50–50 partitions, 60–40 partitions, 70–30 partitions, and 10 – fold cross-validation, the results were 98.86%, 98.46%, and 98.46%, respectively. Wang *et al*. [22] used two common classifiers: SVM and KNN. The suggested feature relevance measurement outperforms previous methods in experiments. It also exceeds many conventional classification execution speed and accuracy methods, demonstrating its usefulness in obtaining the best features. The KNN model had a high accuracy of 95.8%. Oyelade *et al*. [23] used ML methodology to deal with the select and test limitation of one such reasoning algorithm (ST). An efficient input method is an initial step in technique, which allows the system to read, filter, and clean datasets. Knowledge representation frameworks were also built using semantic web languages (ontologies and rule languages) to help the reasoning algorithm. As a result, the ST reasoning structures were modified to support this improvement. The ST-ONCODIAG technique had 81.0% sensitivity and 89.0% specificity.

This project aims to construct ML models and apply them to breast cancer diagnosis. By comparing the results of the eleven classification algorithms and choosing the classification algorithm that achieved the highest results for this dataset, The remainder of this work is divided into the following sections: Section 2 explains the technique, and Section 3 explains the experimental environment. Section 4 explains the results, conclusions, and future work, and Section 5 talks about the significance of the results, conclusions, and future work.

2. RESEARCH METHOD

The major goal of this study is to identify the most accurate and predictive breast cancer detection algorithm. we used different ML classifiers: decision tree, quadratic discriminant analysis, AdaBoost, Bagging meta estimator, extra randomized trees, Gaussian process classifier, Ridge, Gaussian nave Bayes, k-nearest neighbors, multilayer perceptron, and support vector classifier. The first step in our methodology pre-processing includes attribute selection. ML systems that can detect breast cancer using new measures can be improved using pre-processed data. To evaluate the algorithms' performance, we use new data that has been labeled. Our labeled data is often divided into two parts to accomplish this. Splitting the method test train ML models are trained using 80% of the data, known as the training set. We used about 20% of the data, known as the testing data or testing set. When the models have been evaluated, compared the results to see which algorithm is the most accurate and the most probable to diagnose breast cancer.

These algorithms are used for making predictive analysis on ML techniques: A decision tree (DT) is a non-parametric supervised learning approach to classification. The DT is made up of several nodes that

together form a rooted tree, which is a tree with no incoming edges and a "root" node. Each of the other nodes is connected by a single incoming edge. An internal or test node has outer edges; all other nodes are leaves (also known as the decision or terminal nodes. DT is simple to understand and apply. It is possible to imagine trees in your mind. It can work with both categorical and numerical data [24].

Quadratic discriminant analysis (QDA) is an individual covariance matrix is estimated for each class of observations. If you know that each class has a different set of covariances, QDA can be very useful. A problem with QDA is that it can't make things smaller in size. QDA is more flexible than LDA in that it does not require equal variance and covariance. To put it another way, the covariance matrix for each class in QDA might be different. When you have a short training set, LDA is preferable to QDA. QDA has suggested that if the training set is massive and the classifier's variance isn't a key concern, the assumption of a shared covariance matrix for the K classes is unsustainable [25]. AdaBoost (AB) the basic idea behind AB is to fit several weak learners (i.e., models that are just slightly better than random guessings, such as tiny DT) to multiple copies of the data. The final prediction is calculated by averaging the predictions using a weighted majority vote (or mean). AB is a simple program it continually corrects the faults of a poor classifier and enhances accuracy by merging poor learners. You may use AB with several base classifiers. AB does not suffer from overfitting [26].

Bagging Meta-Estimator (BME) is a type of ensemble algorithm that creates many random subsets of the original training data used to test a black-box estimator and combines their different predictions to create the last prediction. These techniques are employed to minimize variation. By adding randomness into the building method of a base estimator (e.g., DT) and then building an ensemble from it. In many situations, bagging approaches provide a straightforward way to enhance one model without changing the algorithm used as a basis. Bagging approaches, in contrast to boosting methods, which are the most effective with weak models, perform best with powerful and complicated models (e.g., fully formed DT) because they give a mechanism to reduce overfitting (e.g., shallow DT) [27]. Extremely randomized trees (ERT) are extremely similar to Random Forests but have two main differences: ERT does not resample data when building a tree, ERT does not use the “best split”. It is usually provided for a small reduction in the model at the cost of a small bias increase [28].

Gaussian process classifier (GPC) more especially for probabilistic classification. The probabilistic classification expresses the test predictions as class probabilities. These binary predictor predictions are combined to provide a multi-class prediction. Gaussian processes have a clear advantage: prediction is probabilistic (Gaussian), allowing actual probability ranges to be GPC offers three major benefits over other widely used classifiers. GPC is capable of dealing with high-dimensional and nonlinear problems that arise in travel mode detection. Instead of determinant classification findings, GPC provides probabilistic outputs, which account for the model uncertainty inherent in trip mode identification. Because GPC is a non-parameterized model, it may tune hyperparameters directly from training data. GPC may also utilize evidence to build a completely automated model selection procedure [29].

Ridge classifier the goal values are transformed to -1, 1 using this classifier. Before treating the problem as a regression problem (multi-output regression in the multiclass case), after which the problem is treated as a regression task, with the same goal. The projected class is determined by the regressor's prediction sign. The problem is addressed as multi-output regression for multiclass classification, and the output with the highest value matches the expected class. Using a (penalized) Least Squares rather than the more common logistic or hinge losses to fit a classification model may appear weird, however, any of those models can provide similar cross-validation scores in terms of or precision/recall accuracy, but the Ridge Classifier's penalized least squares loss allows for a far wider range of numerical solvers with various computational performance profiles [30].

Naïve Bayes (NB) classification algorithms rely on strong assumptions about the independence of variables. NB classifier uses a Gaussian distribution of numeric predictors with mean and standard deviation calculated from the training dataset and assumes independence between predictor variables depending on response. NB models are often employed as a substitute for decision trees to solve classification difficulties. To develop an NB classifier, all rows in the training dataset that include at least one NA will be ignored. Missing values in the test dataset are removed from the probability calculation for making predictions. It performs effectively with categorical input variables compared to numeric input parameters (s). A typical distribution is taken into account (bell curve, which is a powerful inference) [31].

K-nearest neighbors (KNN) in the field of ML and data mining (DM), the KNN classifier is a commonly used and well-known non-parametric classifier that is used to solve a variety of issues. KNN is a simple way to build, but as the dataset grows, the efficiency and speed of the process decrease significantly. KNN performs well with a limited number of input variables, but as the number of input variables increases, it fails to predict the output of more data points. KNN requires rather homogeneous features: If you want to design KNN using a popular length, such as Euclidean or Manhattan distances, absolute differences in

attributes must be weighed the same, i.e., a given distance in feature 1 must signify the same as a given distance in feature 2. Unbalanced data causes complications, and KNN has difficulty dealing with it. KNN is Sensitive to outliers Because it chooses neighbors only based on distance. KNN is fundamentally incapable of dealing with problems requiring partial data. The algorithm is easy to understand and apply. In this case, developing a model, fine-tuning various parameters, or producing more projections are not required [32].

Multi-layer perceptron (MLP) has a hidden layer or layers (except for one input and one output layer). In contrast to a single layer perceptron, MLP is capable of learning functions that are not linear. There are weights connected with all connections, however only three weights (w_0, w_1, and w_2). The input layer is composed of three nodes. The Bias node has a value of 1. Both X_1 and X_2 are used as external inputs by the other two nodes (quantities depending upon the given data). To recap, the input layer does not perform any computation, therefore nodes in the input layer produce anything other than 1, X_1, and the outputs of the hidden layer. The bias node on the hidden layer has an output of 1, making it part of it. The hidden layer's other two nodes' outputs are dependent on the input layer's (1, X_1, X_2) outputs and the connections' weights (edges) [33].

Support vector machine (SVM) is an ML classifier that can generalize across two different classes if it is given a training sample of labeled data. The most important task for the SVM is to find a hyperplane that can distinguish between similar and dissimilar classes of data. Because it works effectively in complex three-dimensional situations, this model is effective. Even when the number of dimensions exceeds the number of samples, the approach remains effective [34].

3. EXPERIMENT ENVIRONMENT

Oyelade *et al*. [23] gathered the WDBC dataset from Madison's University of Wisconsin Hospitals. WDBC is available in the UCI ML repository. To compute features, data from a fine needle aspirate (FNA) of a breast lump is employed; cell nuclei are defined using the image's features. The WDBC dataset contained 569 patients (62.74 % benign, 37.26 % malignant) with WDBC as the patient ID, 30 tumor features, and one class indicator [35]. In total, 30 parameters were considered: area, texture, radius, perimeter, compactness, smoothness, concavity, concave points, and symmetry. The ID number was removed because it has no bearing on the classification process. For all of the experiments on the ML algorithms (DT, QDA, AB, BME, ERT, GPC, Ridge, NB, KNN, MLP, and SVC) presented in this paper, the Scikit-Learn library and the Python language were utilized. It was implemented with Python's SciPy, pandas, and matplotlib libraries.

4. RESULTS AND DISCUSSION

Eleven classification methods were used to divide WDBC into benign and malignant tumors: DT, QDA, AB, BME, ERT, GPC, Ridge, NB, KNN, MLP, and SVC, to assess each model, we used the set of criteria: specificity sensitivity precision accuracy, F1-Score. Interestingly, it can note these algorithms, GPC, Ridge, KNN, and SVC have a maximum specificity of 100%, DT has the lowest specificity with a value of 89.55%. The results of specificity are presented in Figure 1.

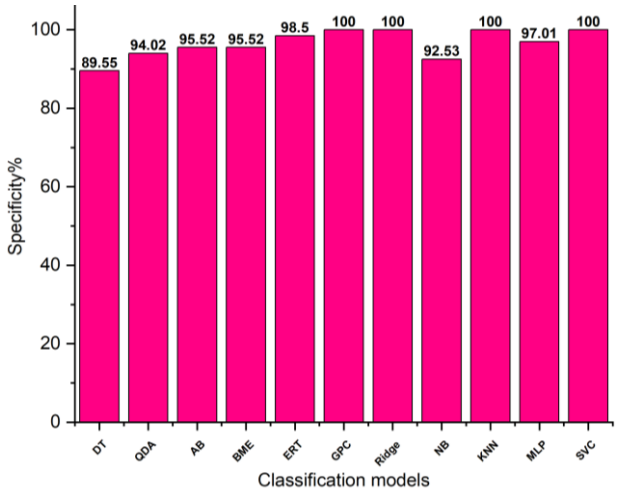

Figure 1. The specificity of classification models

Figure 2 illustrated that QDA, AB, BME, ERT, and MLP had the maximum sensitivity of 95.74%, but GPC, Ridge, Gaussian NB, and KNN had the lowest sensitivity 89.36%. The results indicate the highest accuracy was 97.36% in ERT, but, 91.22% in DT and Gaussian NB, as shown in Figure 3.

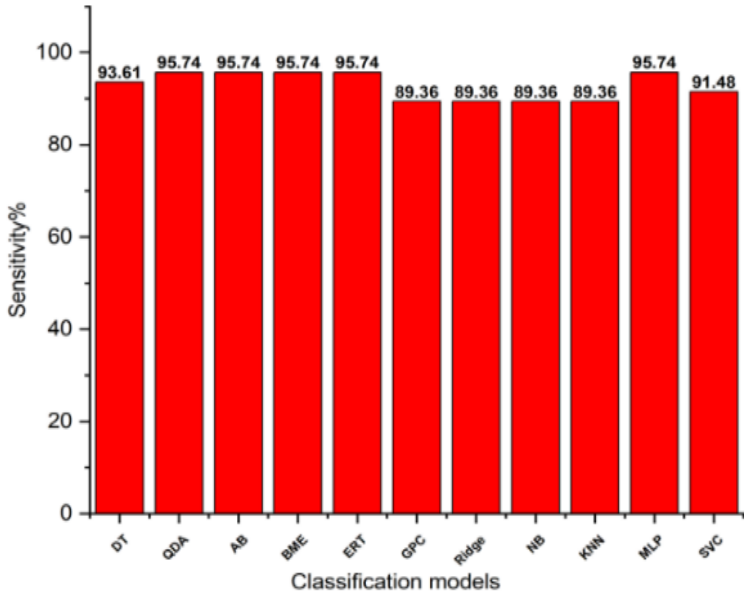

Figure 2. The sensitivity of classification models

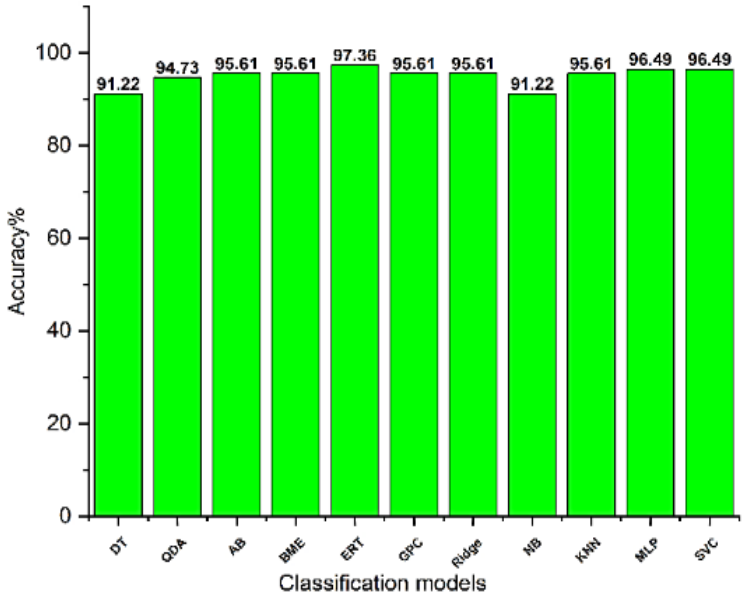

Figure 3. The accuracy of classification models

It is apparent that the GPC, Ridge, KNN, and SVC all had a precision of 100%, while DT had a precision of 86.27% as shown in Figure 4. The ERT classifier scored the highest F1-Score of 96.77%, while the Gaussian Naive Bayes classifier had the lowest F1-Score of 89.36%, as illustrated in Figure 5.

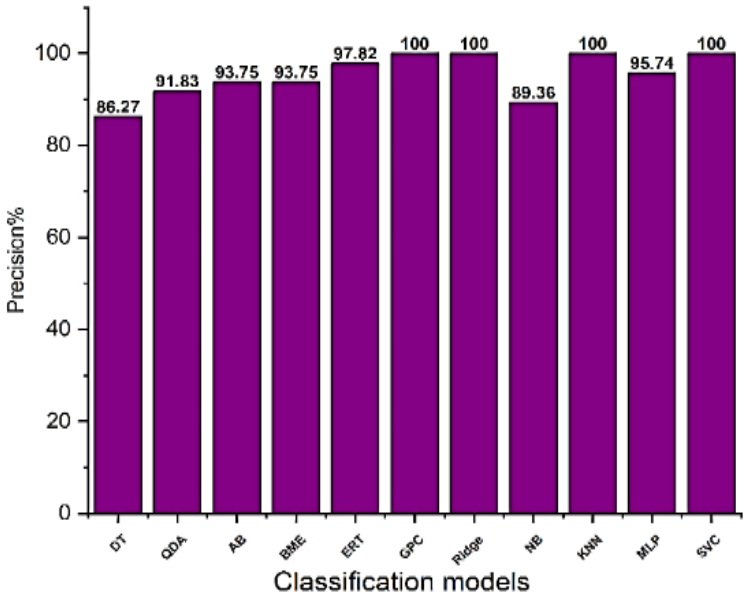

Figure 4. The precision of classification models

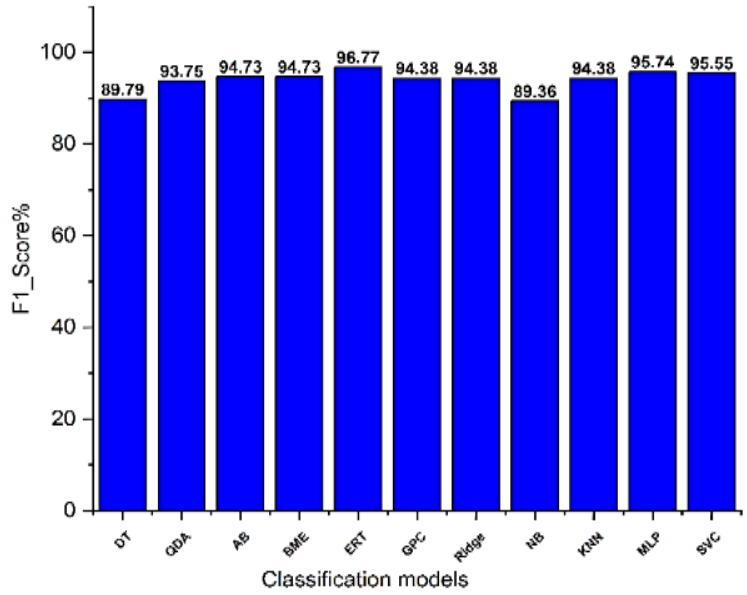

Figure 5. F1_score for classification models

The most remarkable result from this study is the performance of ML in the prediction and detection of breast cancer. The mean accuracy of all applied models was 95.05%. We observed that the ERT model made the highest results with this data, was stable, and produced superior results (F1- score in especially). Additionally, clinicians are interested in the results of the sensitivity-specificity trade-off. Therefore, sensitivity is a term that relates to a diagnostic test's capacity to diagnose cancer patients as abnormal.

In comparison, specificity refers to a diagnostic's ability to designate non-cancerous patients as normal. For ML programmers, the receiver operating characteristic (ROC) curve results are of primary importance. The receiver operating characteristic ROC curve is constructed by computing and plotting the true positive rate vs. the false positive rate. Because it demonstrates the model's stability and reliability, it attained a maximum of 1.00 in ERT, as shown in Figure 6.

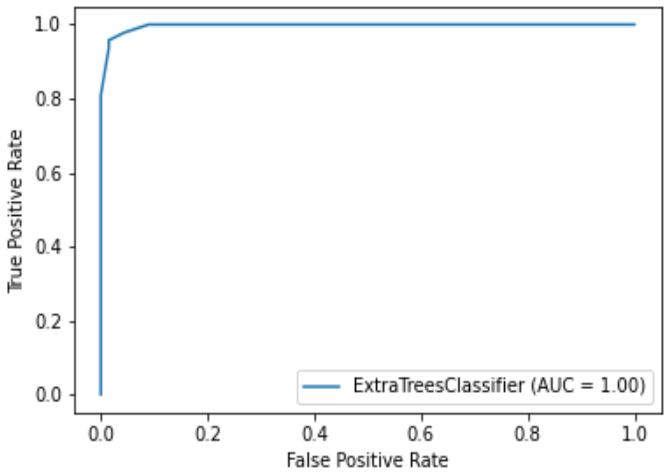

Figure 6. ROC curve for ERT

This outcome can be explained because ERT is an ensemble method that learns from previous predictor mistakes to improve future predictions. The strategy combines the weaker classification model with a powerful learner, increasing the model's predictability. By sequentially arranging weak learners, weak learners can learn from previous ones, strengthening prediction models. Ensemble approaches are well suited for reducing model variability and increasing prediction accuracy. When multiple models are combined into a single prediction, the variance is reduced. To correct for previous prediction errors, new predictors are fit. By adding predictors consecutively to the ensemble, ERT improves the model's accuracy. It makes use of the gradient to identify and rectify mistakes in the predictions made by learners. Table 1 compares the most significant findings obtained in this study to other outcomes presented in the literature. These results are consistent with those of previous studies that show ML can accurately predict the presence of breast cancer. To summarize, the ML models achieved a greater classification accuracy rate, reduced false positives, and improved performance. As a result of this research, it has been demonstrated that the ML can help a radiologist make an accurate breast cancer diagnosis and that it can demonstrate significant capability in the domain of medical decision-making.

Table1. Comparison of performance to previous studies results

Reference	Year	ML technique	Accuracy	Sensitivity	Specificity	Precision	AUC
Salama *et al.* [18]	2012	MLP	97.71%	-	-	-	-
Azar and El-Said [19]	2013	ANN	97.66%	98.65%	95.82%	97.77%	0.993
Senapati *et al.* [20]	2013	Local linear wavelet neural networks	97.20%	-	-	-	-
Dora *et al.* [21]	2017	GNRBA	98.25%	-	-	-	-
Wang and Feng [22]	2018	KNN	95.80%	-	-	-	-
Oyelade *et al.* [23]	2018	ST-ONCODIAG	0.81%	-	89.00%	-	-
Our work		ERT	97.36%	95.74%	98.50%	97.82%	1.00

5. CONCLUSION

One of the most critical scientific subjects being pursued is the early detection of breast cancer. It is a terrible disease that affects women all over the world. In women, it is one of the most frequent cancer types. One in every eight women in the world is at risk of being diagnosed at a certain point in their lives. Accurate classification of breast cancer tumors has become a complex problem in the medical world. The study aims to compare the performance of different classification methods. The performance of ML in breast cancer prediction and diagnosis has been described as the most surprising result of these results. As evidenced by this study, the ERT classifier surpasses the WDBC in terms of specificity and precision. We estimate that the high results will be attributed to the fact that using ensembles rather than a single predictive model improves predictive modeling performance in this scenario, as this model belongs to the ensemble methods group. These limitations will require additional research in the future. As a result, future work should evaluate machine learning models for new medical diagnostic challenges and optimize their performance using high-performance computing methods.

ACKNOWLEDGEMENTS

The authors would like to thank Mustansiriya University for their valuable support and providing the essential facilities for this research.

REFERENCES

[1] M. Y. Kamil, "Computer-aided diagnosis system for breast cancer based on the Gabor filter technique," *International Journal of Electrical Computer Engineering*, vol. 10, no. 5, pp. 5235-5242, 2020, doi: 10.11591/ijece.v10i5.pp5235-5242.

[2] P. N. Pandey, N. Saini, N. Sapre, A. Kulkarni, and A. K. Tiwari, "Prioritising breast cancer theranostics: A current medical longing in oncology," *Cancer Treatment and Research Communications*, vol. 29, p. 100465, 2021, doi: 10.1016/j.ctarc.2021.100465.

[3] A. Conti, A. Duggento, I. Indovina, M. Guerrisi, and N. Toschi, "Radiomics in breast cancer classification and prediction," *Seminars in Cancer Biology*, vol. 72, pp. 238-250, 2021, doi: 10.1016/j.semcancer.2020.04.002.

[4] S. Khairnar, S. D. Thepade, and S. Gite, "Effect of image binarization thresholds on breast cancer identification in mammography images using OTSU, Niblack, Burnsen, Thepade's SBTC," *Intelligent Systems with Applications*, vol. 10-11, p. 200046, 2021, doi: 10.1016/j.iswa.2021.200046.

[5] E. A. Radhi and M. Y. Kamil, "Breast tumor detection via active contour technique," *International Journal of Intelligent Engineering and Systems*, vol. 14, no. 4, pp. 561-570, 2021, doi: 10.22266/ijies2021.0831.49.

[6] M. Y. Kamil and E. A. Radhi, "Breast tumor segmentation in mammography image via Chan-Vese technique," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 22, no. 2, pp. 201-209, 2021, doi: 10.11591/ijeecs.v22.i2.pp809-817.

[7] R. Vijayarajeswari, P. Parthasarathy, S. Vivekanandan, and A. A. Basha, "Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform," *Measurement*, vol. 146, pp. 800-805, 2019, doi: 10.1016/j.measurement.2019.05.083.

[8] S. Halvaei *et al*., "Exosomes in Cancer liquid biopsy: a focus on breast cancer," *Molecular Therapy - Nucleic Acids*, vol. 10, pp. 131-141, 2018, doi: 10.1016/j.omtn.2017.11.014.

[9] I. Steinberg, D. M. Huland, O. Vermesh, H. E. Frostig, W. S. Tummers, and S. S. Gambhir, "Photoacoustic clinical imaging," *Photoacoustics*, vol. 14, pp. 77-98, 2019, doi: 10.1016/j.pacs.2019.05.001.

[10] N. I. R. Yassin, S. Omran, E. M. F. El Houby, and H. Allam, "Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: A systematic review," *Computer Methods and Programs in Biomedicine*, vol. 156, pp. 25-45, 2018, doi: 10.1016/j.cmpb.2017.12.012.

[11] S. Pandey, A. Sharma, M. K. Siddiqui, D. Singla, and J. Vanderpuye-Orgle, "ai3 prediction of breast cancer using k-nearest neighbour: a supervised machine learning algorithm," *Value in Health*, vol. 23, p. S1, 2020, doi: 10.1016/j.jval.2020.04.006.

[12] C. Kaushal, S. Bhat, D. Koundal, and A. Singla, "Recent trends in computer assisted diagnosis (CAD) system for breast cancer diagnosis using histopathological images," *IRBM*, vol. 40, no. 4, pp. 211-227, 2019, doi: 10.1016/j.irbm.2019.06.001.

[13] A. H. Farhan and M. Y. Kamil, "Texture analysis of breast cancer via LBP, HOG, and GLCM techniques," in *IOP Conference Series: Materials Science and Engineering*, vol. 928, no. 7, p. 072098, 2020, doi: 10.1088/1757-899X/928/7/072098.

[14] A. H. Farhan and M. Y. Kamil, "Texture analysis of mammogram using local binary pattern method," in *Journal of Physics: Conference Series*, vol. 1530, no. 1, p. 012091, 2020, doi: 10.1088/1742-6596/1530/1/012091.

[15] M. Y. Kamil and A.-L. A. Jassam, "Analysis of tissue abnormality in mammography images using gray level co-occurrence matrix method," in *Journal of Physics: Conference Series*, vol. 1530, no. 1, p. 012101, 2020, doi: 10.1088/1742-6596/1530/1/012101.

[16] M. Tariq, S. Iqbal, H. Ayesha, I. Abbas, K. T. Ahmad, and M. F. K. Niazi, "Medical image based breast cancer diagnosis: State of the art and future directions," *Expert Systems with Applications*, vol. 167, p. 114095, 2021, doi: 10.1016/j.eswa.2020.114095.

[17] S. Saxena and M. Gyanchandani, "Machine learning methods for computer-aided breast cancer diagnosis using histopathology: a narrative review," *Journal of Medical Imaging and Radiation Sciences*, vol. 51, no. 1, pp. 182-193, 2020, doi: 10.1016/j.jmir.2019.11.001.

[18] G. I. Salama, M. Abdelhalim, and M. A. Zeid, "Breast cancer diagnosis on three different datasets using multi-classifiers," *International Journal of Computer and Information Technology*, vol. 1, no. 1, p. 2, 2012.

[19] A. T. Azar and S. A. El-Said, "Probabilistic neural network for breast cancer classification," *Neural Computing and Applications*, vol. 23, no. 6, pp. 1737-1751, 2013, doi: 10.1007/s00521-012-1134-8.

[20] M. R. Senapati, A. K. Mohanty, S. Dash, and P. K. Dash, "Local linear wavelet neural network for breast cancer recognition," *Neural Computing and Applications*, vol. 22, no. 1, pp. 125-131, 2013, doi: 10.1007/s00521-011-0670-y.

[21] L. Dora, S. Agrawal, R. Panda, and A. Abraham, "Optimal breast cancer classification using Gauss–Newton representation based algorithm," *Expert Systems with Applications*, vol. 85, pp. 134-145, 2017, doi: 10.1016/j.eswa.2017.05.035.

[22] Y. Wang and L. Feng, "Hybrid feature selection using component co-occurrence based feature relevance measurement," *Expert Systems with Applications*, vol. 102, pp. 83-99, 2018, doi: 10.1016/j.eswa.2018.01.041.

[23] O. N. Oyelade, A. A. Obiniyi, S. B. Junaidu, and S. A. Adewuyi, "ST-ONCODIAG: A semantic rule-base approach to diagnosing breast cancer base on Wisconsin datasets," *Informatics in Medicine Unlocked*, vol. 10, pp. 117-125, 2018, doi: 10.1016/j.imu.2017.12.008.

[24] A. T. Azar and S. M. El-Metwally, "Decision tree classifiers for automated medical diagnosis," *Neural Computing and Applications*, vol. 23, no. 7, pp. 2387-2403, 2013, doi: 10.1007/s00521-012-1196-7.

[25] U. Raghavendra, U. Rajendra Acharya, H. Fujita, A. Gudigar, J. H. Tan, and S. Chokkadi, "Application of Gabor wavelet and Locality Sensitive Discriminant Analysis for automated identification of breast cancer using digitized mammogram images," *Applied Soft Computing*, vol. 46, pp. 151-161, 2016, doi: 10.1016/j.asoc.2016.04.036.

[26] B. Bigdeli, P. Pahlavani, and H. A. Amirkolaee, "An ensemble deep learning method as data fusion system for remote sensing multisensor classification," *Applied Soft Computing*, vol. 110, p. 107563, 2021, doi: 10.1016/j.asoc.2021.107563.

[27] S. Moral-García, C. J. Mantas, J. G. Castellano, M. D. Benítez, and J. Abellán, "Bagging of credal decision trees for imprecise classification," *Expert Systems with Applications*, vol. 141, p. 112944, 2020, doi: 10.1016/j.eswa.2019.112944.

[28] Y. Liang, S. Zhang, H. Qiao, and Y. Yao, "iPromoter-ET: Identifying promoters and their strength by extremely randomized trees-based feature selection," *Analytical Biochemistry*, vol. 630, p. 114335, 2021, doi: 10.1016/j.ab.2021.114335.

[29] H. Oyama, M. Yamakita, K. Sata, and A. Ohata, "Identification of static boundary model based on Gaussian Process classification," *IFAC-PapersOnLine*, vol. 49, no. 11, pp. 787-792, 2016, doi: 10.1016/j.ifacol.2016.08.115.

[30] X. Zhang, W. Chao, Z. Li, C. Liu, and R. Li, "Multi-modal kernel ridge regression for social image classification," *Applied Soft Computing*, vol. 67, pp. 117-125, 2018, doi: 10.1016/j.asoc.2018.02.030.
[31] R. Blanquero, E. Carrizosa, P. Ramírez-Cobo, and M. R. Sillero-Denamiel, "Variable selection for Naïve Bayes classification," *Computers & Operations Research*, vol. 135, p. 105456, 2021, doi: 10.1016/j.cor.2021.105456.
[32] C. Viji, J. Beschi Raja, R. S. Ponmagal, S. T. Suganthi, P. Parthasarathi, and S. Pandiyan, "Efficient fuzzy based K-nearest neighbour technique for web services classification," *Microprocessors and Microsystems*, vol. 76, p. 103097, 2020, doi: 10.1016/j.micpro.2020.103097.
[33] S. Radhakrishnan, S. G. Nair, and J. Isaac, "Multilayer perceptron neural network model development for mechanical ventilator parameters prediction by real time system learning," *Biomedical Signal Processing and Control*, vol. 71, p. 103170, 2022, doi: 10.1016/j.bspc.2021.103170.
[34] B. A. Akinnuwesi, B. O. Macaulay, and B. S. Aribisala, "Breast cancer risk assessment and early diagnosis using principal component analysis and support vector machine techniques," *Informatics in Medicine Unlocked*, vol. 21, p. 100459, 2020, doi: 10.1016/j.imu.2020.100459.
[35] M. H. Alshayeji, H. Ellethy, S. Abed, and R. Gupta, "Computer-aided detection of breast cancer on the Wisconsin dataset: An artificial neural networks approach," *Biomedical Signal Processing and Control*, vol. 71, p. 103141, 2022, doi: 10.1016/j.bspc.2021.103141.

BIOGRAPHIES OF AUTHORS

Rania R. Kadhim holds a Master of Science (M.Sc.) in physics from Mustansiriya University, College of Science, in 2016. She is currently a Ph.D. student in digital image processing at Mustansiriya University in Iraq. Her research areas of interest include medical image processing, machine learning, computer vision, and image classification. She can be contacted at email: r90r61@gmail.com.

Mohammed Y. Kamil holds a Master of Science (M.Sc.) in Optics from Mustansiriya University, in 2005. and Ph.D. in digital image processing from Mustansiriya University, Iraq, in 2011, besides several professional certificates and skills. He is currently a professor with the physics department at Mustansiriya University, Baghdad, Iraq. He is a member of the Institute of Electrical and Electronics Engineers (IEEE) and an editor in the Iraqi Journal of Physics. His research areas of interest include medical image processing, computer vision, and Artificial Intelligence. He can be contacted at email: m80y98@uomustansiriyah.edu.iq.